

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 101612,466B  
Source: IFWO  
Date Processed by STIC: 1-31-05

***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:14

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw

```

3 <110> APPLICANT: Madison, Edwin
4     Ong, Edgar
5     Yeh, Juinn-Chern
7 <120> TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE 16, THE
8     ENCODED PROTEINS AND METHODS BASED THEREON
10 <130> FILE REFERENCE: 24745-1625
12 <140> CURRENT APPLICATION NUMBER: 10/612,466B
13 <141> CURRENT FILING DATE: 2003-07-01
15 <150> PRIOR APPLICATION NUMBER: 60/394,347
16 <151> PRIOR FILING DATE: 2002-07-02
18 <160> NUMBER OF SEQ ID NOS: 22
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 3147
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo Sapien
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (23)...(2589)
30 <223> OTHER INFORMATION: Nucleotide sequence encoding MTSP1
32 <300> PUBLICATION INFORMATION:
33 <308> DATABASE ACCESSION NO: GenBank #AR081724
34 <309> DATABASE ENTRY DATE: 2000-08-31
36 <400> SEQUENCE: 1
37 tcaagagcgg cctcggggta cc atg ggg agc gat cgg gcc cgc aag ggc gga 52
38                               Met Gly Ser Asp Arg Ala Arg Lys Gly Gly
39                               1           5           10
41 ggg ggc ccg aag gac ttc ggc gcg gga ctc aag tac aac tcc cgg cac 100
42 Gly Gly Pro Lys Asp Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His
43           15           20           25
45 gag aaa gtg aat ggc ttg gag gaa ggc gtg gag ttc ctg cca gtc aac 148
46 Glu Lys Val Asn Gly Leu Glu Glu Gly Val Glu Phe Leu Pro Val Asn
47           30           35           40
49 aac gtc aag aag gtg gaa aag cat ggc ccg ggg cgc tgg gtg gtg ctg 196
50 Asn Val Lys Lys Val Glu Lys His Gly Pro Gly Arg Trp Val Val Leu
51           45           50           55
53 gca gcc gtg ctg atc ggc ctc ctc ttg gtc ttg ctg ggg atc ggc ttc 244
54 Ala Ala Val Leu Ile Gly Leu Leu Leu Val Leu Leu Gly Ile Gly Phe
55           60           65           70
57 ctg gtg tgg cat ttg cag tac cgg gac gtg cgt gtc cag aag gtc ttc 292
58 Leu Val Trp His Leu Gln Tyr Arg Asp Val Arg Val Gln Lys Val Phe
59 75           80           85           90
61 aat ggc tac atg agg atc aca aat gag aat ttt gtg gat gcc tac gag 340

```

## RAW SEQUENCE LISTING

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:14

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw

62	Asn	Gly	Tyr	Met	Arg	Ile	Thr	Asn	Glu	Asn	Phe	Val	Asp	Ala	Tyr	Glu	
63					95					100					105		
65	aac	tcc	aac	tcc	act	gag	ttt	gta	agc	ctg	gcc	agc	aag	gtg	aag	gac	388
66	Asn	Ser	Asn	Ser	Thr	Glu	Phe	Val	Ser	Leu	Ala	Ser	Lys	Val	Lys	Asp	
67				110						115				120			
69	gcg	ctg	aag	ctg	ctg	tac	agc	gga	gtc	cca	ttc	ctg	ggc	ccc	tac	cac	436
70	Ala	Leu	Lys	Leu	Leu	Tyr	Ser	Gly	Val	Pro	Phe	Leu	Gly	Pro	Tyr	His	
71			125					130					135				
73	aag	gag	tcg	gct	gtg	acg	gcc	ttc	agc	gag	ggc	agc	gtc	atc	gcc	tac	484
74	Lys	Glu	Ser	Ala	Val	Thr	Ala	Phe	Ser	Glu	Gly	Ser	Val	Ile	Ala	Tyr	
75		140					145				150						
77	tac	tggt	tct	gag	ttc	agc	atc	ccg	cag	cac	ctg	gtg	gag	gag	gcc	gag	532
78	Tyr	Trp	Ser	Glu	Phe	Ser	Ile	Pro	Gln	His	Leu	Val	Glu	Glu	Ala	Glu	
79	155				160					165				170			
81	cgc	gtc	atg	gcc	gag	gag	cgc	gta	gtc	atg	ctg	ccc	ccg	cgg	gcg	cgc	580
82	Arg	Val	Met	Ala	Glu	Arg	Val	Val	Met	Leu	Pro	Pro	Arg	Ala	Arg		
83				175				180					185				
85	tcc	ctg	aag	tcc	ttt	gtg	gtc	acc	tca	gtg	gtg	gct	ttc	ccc	acg	gac	628
86	Ser	Leu	Lys	Ser	Phe	Val	Val	Thr	Ser	Val	Val	Ala	Phe	Pro	Thr	Asp	
87			190					195					200				
89	tcc	aaa	aca	gta	cag	agg	acc	cag	gac	aac	agc	tgc	agc	ttt	ggc	ctg	676
90	Ser	Lys	Thr	Val	Gln	Arg	Thr	Gln	Asp	Asn	Ser	Cys	Ser	Phe	Gly	Leu	
91		205					210				215						
93	cac	gcc	cgc	ggg	gtg	gag	ctg	atg	cgc	ttc	acc	acg	ccc	ggc	ttc	cct	724
94	His	Ala	Arg	Gly	Val	Glu	Leu	Met	Arg	Phe	Thr	Thr	Pro	Gly	Phe	Pro	
95		220				225				230							
97	gac	agc	ccc	tac	ccc	gct	cat	gcc	cgc	tgc	cag	tggt	gcc	ctg	cgg	ggg	772
98	Asp	Ser	Pro	Tyr	Pro	Ala	His	Ala	Arg	Cys	Gln	Trp	Ala	Leu	Arg	Gly	
99	235			240				245				250					
101	gac	gcc	gac	tca	gtg	ctg	agc	ctc	acc	ttc	cgc	agc	ttt	gac	ctt	gcg	820
102	Asp	Ala	Asp	Ser	Val	Leu	Ser	Leu	Thr	Phe	Arg	Ser	Phe	Asp	Leu	Ala	
103				255				260					265				
105	tcc	tgc	gac	gag	cgc	ggc	agc	gac	ctg	gtg	acg	gtg	tac	aac	acc	ctg	868
106	Ser	Cys	Asp	Glu	Arg	Gly	Ser	Asp	Leu	Val	Thr	Val	Tyr	Asn	Thr	Leu	
107			270					275					280				
109	agc	ccc	atg	gag	ccc	cac	gcc	ctg	gtg	cag	ttg	tgt	ggc	acc	tac	cct	916
110	Ser	Pro	Met	Glu	Pro	His	Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	
111		285				290				295							
113	ccc	tcc	tac	aac	ctg	acc	ttc	cac	tcc	tcc	cag	aac	gtc	ctg	ctc	atc	964
114	Pro	Ser	Tyr	Asn	Leu	Thr	Phe	His	Ser	Ser	Gln	Asn	Val	Leu	Leu	Ile	
115		300				305				310							
117	aca	ctg	ata	acc	aac	act	gag	cgg	cgg	cat	ccc	ggc	ttt	gag	gcc	acc	1012
118	Thr	Leu	Ile	Thr	Asn	Thr	Glu	Arg	Arg	His	Pro	Gly	Phe	Glu	Ala	Thr	
119	315				320				325				330				
121	ttc	ttc	cag	ctg	cct	agg	atg	agc	agc	tgt	gga	ggc	cgc	tta	cgt	aaa	1060
122	Phe	Phe	Gln	Leu	Pro	Arg	Met	Ser	Ser	Cys	Gly	Gly	Arg	Leu	Arg	Lys	
123				335				340					345				
125	gcc	cag	ggg	aca	ttc	aac	agc	ccc	tac	tac	cca	ggc	cac	tac	cca	ccc	1108
126	Ala	Gln	Gly	Thr	Phe	Asn	Ser	Pro	Tyr	Tyr	Pro	Gly	His	Tyr	Pro	Pro	

## RAW SEQUENCE LISTING

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:14

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw

127		350		355		360		
129	aac att gac tgc aca tgg aac att gag gtg ccc aac aac cag cat gtg	1156						
130	Asn Ile Asp Cys Thr Trp Asn Ile Glu Val Pro Asn Asn Gln His Val							
131		365		370		375		
133	aag gtg agc ttc aaa ttc ttc tac ctg ctg gag ccc ggc gtg cct gcg	1204						
134	Lys Val Ser Phe Lys Phe Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala							
135		380		385		390		
137	ggc acc tgc ccc aag gac tac gtg gag atc aat ggg gag aaa tac tgc	1252						
138	Gly Thr Cys Pro Lys Asp Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys							
139	395		400		405		410	
141	gga gag agg tcc cag ttc gtc gtc acc agc aac agc aac aag atc aca	1300						
142	Gly Glu Arg Ser Gln Phe Val Val Thr Ser Asn Ser Asn Lys Ile Thr							
143		415		420		425		
145	gtt cgc ttc cac tca gat cag tcc tac acc gac acc ggc ttc tta gct	1348						
146	Val Arg Phe His Ser Asp Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala							
147		430		435		440		
149	gaa tac ctc tcc tac gac tcc agt gac cca tgc ccg ggg cag ttc acg	1396						
150	Glu Tyr Leu Ser Tyr Asp Ser Ser Asp Pro Cys Pro Gly Gln Phe Thr							
151		445		450		455		
153	tgc cgc acg ggg cgg tgt atc cgg aag gag ctg cgc tgt gat ggc tgg	1444						
154	Cys Arg Thr Gly Arg Cys Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp							
155		460		465		470		
157	gcc gac tgc acc gac cac agc gat gag ctc aac tgc agt tgc gac gcc	1492						
158	Ala Asp Cys Thr Asp His Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala							
159	475		480		485		490	
161	ggc cac cag ttc acg tgc aag aac aag ttc tgc aag ccc ctc ttc tgg	1540						
162	Gly His Gln Phe Thr Cys Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp							
163		495		500		505		
165	gtc tgc gac agt gtg aac gac tgc gga gac aac agc gac gag cag ggg	1588						
166	Val Cys Asp Ser Val Asn Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly							
167		510		515		520		
169	tgc agt tgt ccg gcc cag acc ttc agg tgt tcc aat ggg aag tgc ctc	1636						
170	Cys Ser Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu							
171		525		530		535		
173	tcg aaa agc cag cag tgc aat ggg aag gac gac tgt ggg gac ggg tcc	1684						
174	Ser Lys Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser							
175		540		545		550		
177	gac gag gcc tcc tgc ccc aag gtg aac gtc gtc act tgt acc aaa cac	1732						
178	Asp Glu Ala Ser Cys Pro Lys Val Asn Val Val Thr Cys Thr Lys His							
179	555		560		565		570	
181	acc tac cgc tgc ctc aat ggg ctc tgc ttg agc aag ggc aac cct gag	1780						
182	Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu							
183		575		580		585		
185	tgt gac ggg aag gag gac tgt agc gac ggc tca gat gag aag gac tgc	1828						
186	Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp Glu Lys Asp Cys							
187		590		595		600		
189	gac tgt ggg ctg cgg tca ttc acg aga cag gct cgt gtt gtt ggg ggc	1876						
190	Asp Cys Gly Leu Arg Ser Phe Thr Arg Gln Ala Arg Val Val Gly Gly							
191		605		610		615		

## RAW SEQUENCE LISTING

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:14

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw

```

193 acg gat gcg gat gag ggc gag tgg ccc tgg cag gta agc ctg cat gct 1924
194 Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln Val Ser Leu His Ala
195      620      625      630
197 ctg ggc cag ggc cac atc tgc ggt gct tcc ctc atc tct ccc aac tgg 1972
198 Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp
199 635      640      645      650
201 ctg gtc tct gcc gca cac tgc tac atc gat gac aga gga ttc agg tac 2020
202 Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp Arg Gly Phe Arg Tyr
203      655      660      665
205 tca gac ccc acg cag tgg acg gcc ttc ctg ggc ttg cac gac cag agc 2068
206 Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly Leu His Asp Gln Ser
207      670      675      680
209 cag cgc agc gcc cct ggg gtg cag gag cgc agg ctc aag cgc atc atc 2116
210 Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg Leu Lys Arg Ile Ile
211      685      690      695
213 tcc cac ccc ttc ttc aat gac ttc acc ttc gac tat gac atc gcg ctg 2164
214 Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu
215      700      705      710
217 ctg gag ctg gag aaa ccg gca gag tac agc tcc atg gtg cgg ccc atc 2212
218 Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser Met Val Arg Pro Ile
219 715      720      725      730
221 tgc ctg ccg gac gcc tcc cat gtc ttc cct gcc ggc aag gcc atc tgg 2260
222 Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala Gly Lys Ala Ile Trp
223      735      740      745
225 gtc acg ggc tgg gga cac acc cag tat gga ggc act ggc gcg ctg atc 2308
226 Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile
227      750      755      760
229 ctg caa aag ggt gag atc cgc gtc atc aac cag acc acc tgc gag aac 2356
230 Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln Thr Thr Cys Glu Asn
231      765      770      775
233 ctc ctg ccg cag cag atc acg ccg cgc atg atg tgc gtg ggc ttc ctc 2404
234 Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met Cys Val Gly Phe Leu
235      780      785      790
237 agc ggc ggc gtg gac tcc tgc cag ggt gat tcc ggg gga ccc ctg tcc 2452
238 Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser
239 795      800      805      810
241 agc gtg gag gcg gat ggg cgg atc ttc cag gcc ggt gtg gtg agc tgg 2500
242 Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly Val Val Ser Trp
243      815      820      825
245 gga gac ggc tgc gct cag agg aac aag cca ggc gtg tac aca agg ctc 2548
246 Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu
247      830      835      840
249 cct ctg ttt cgg gac tgg atc aaa gag aac act ggg gta ta ggggcccgggg 2599
250 Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly Val
251      845      850      855
253 ccacccaaat gtgtacacct gcgggggccac ccategtcca cccagtggtg cagcctgca 2659
254 ggctggagac tggaccgctg actgcaccag cgccccaga acatacactg tgaactcaat 2719
255 ctccagggtc ccaaattctgc ctagaaaacc tctcgtttcc tcagcctcca aagtggagct 2779
256 gggaggtaga aggggaggac actggtggtt ctactgaccc aactgggggc aaaggtttga 2839

```

## RAW SEQUENCE LISTING

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:14

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw

```

257 agacacagcc tcccccgcca gcccgaagct gggccgaggc gcgtttgtgt atatctgcct 2899
258 cccctgtctg taaggagcag cgggaacgga gcttcggagc ctctcagtg aaggtggtgg 2959
259 ggctgccgga tctgggctgt ggggcccttg ggcacgctc ttgaggaagc ccaggctcgg 3019
260 aggaccctgg aaaacagacg ggtctgagac tgaaattgtt ttaccagctc ccagggtgga 3079
261 cttcagtgtg tgtatttgtg taaatgggta aaacaattta tttcttttta aaaaaaaaaa 3139
262 aaaaaaaaaa 3147
264 <210> SEQ ID NO: 2
265 <211> LENGTH: 855
266 <212> TYPE: PRT
267 <213> ORGANISM: Homo Sapien
269 <400> SEQUENCE: 2
270 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Gly Pro Lys Asp Phe
271 1 5 10 15
272 Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn Gly Leu
273 20 25 30
274 Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys Lys Val Glu
275 35 40 45
276 Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala Val Leu Ile Gly
277 50 55 60
278 Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln
279 65 70 75 80
280 Tyr Arg Asp Val Arg Val Gln Lys Val Phe Asn Gly Tyr Met Arg Ile
281 85 90 95
282 Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu Asn Ser Asn Ser Thr Glu
283 100 105 110
284 Phe Val Ser Leu Ala Ser Lys Val Lys Asp Ala Leu Lys Leu Leu Tyr
285 115 120 125
286 Ser Gly Val Pro Phe Leu Gly Pro Tyr His Lys Glu Ser Ala Val Thr
287 130 135 140
288 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser
289 145 150 155 160
290 Ile Pro Gln His Leu Val Glu Glu Ala Glu Arg Val Met Ala Glu Glu
291 165 170 175
292 Arg Val Val Met Leu Pro Pro Arg Ala Arg Ser Leu Lys Ser Phe Val
293 180 185 190
294 Val Thr Ser Val Val Ala Phe Pro Thr Asp Ser Lys Thr Val Gln Arg
295 195 200 205
296 Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu His Ala Arg Gly Val Glu
297 210 215 220
298 Leu Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala
299 225 230 235 240
300 His Ala Arg Cys Gln Trp Ala Leu Arg Gly Asp Ala Asp Ser Val Leu
301 245 250 255
302 Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala Ser Cys Asp Glu Arg Gly
303 260 265 270
304 Ser Asp Leu Val Thr Val Tyr Asn Thr Leu Ser Pro Met Glu Pro His
305 275 280 285
306 Ala Leu Val Gln Leu Cys Gly Thr Tyr Pro Pro Ser Tyr Asn Leu Thr
307 290 295 300

```

**VERIFICATION SUMMARY**

DATE: 01/31/2005

PATENT APPLICATION: US/10/612,466B

TIME: 11:50:15

Input Set : A:\1625seq.003

Output Set: N:\CRF4\01312005\J612466B.raw